

## Final Report to

Office of Naval Research  
Code 252B:JGW  
Ballston Tower One  
800 North Quincy Street  
Arlington, Virginia 22217-5660  
(CFDA No.: 12.300)

Title: Dialogue Theory for Virtual Environments  
Grant No. N00014-94-1-0938  
R&T Project:3331005vei01

Principal Investigator: Alan W. Biermann  
(Phone 919-660-6500; email awb@cs.duke.edu)

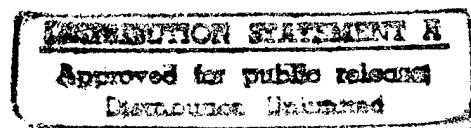
Institution:  
Duke University  
Durham, North Carolina 27708

30 November 1996

### Table of Contents:

1. Principal Investigator.....	2
2. Productivity Measures.....	2
3. Summary of Objectives and Approach.....	2
4. Detailed Summary of Technical Progress.....	3
5. Transitions and DOD Interactions.....	3
6. Software and Hardware Prototypes.....	4
7. List of Publications.....	5
8. Invited and Contributed Presentations.....	5
9. Honors, Prizes or Awards Received.....	6
10. Project Personnel Promotions.....	6
11. Project Staff.....	6
12. Multimedia URL.....	6
13. Keywords.....	6
14. Business Office.....	7
15. Expenditures.....	7
16. Current and Former Students.....	7
17. Book Plans.....	8
18. Sabbatical Plans.....	8
19. Related Research.....	8
20. History.....	8

19970325 001



## **1. Principal Investigator.**

- \* PI Name: Dr. Alan W. Biermann
  - \* PI Institution: Duke University
  - \* PI Phone Number: 919.660.6500
  - \* PI Fax Number: 919.660.6519
  - \* PI Street Address: Computer Science Department, Box 90129
  - \* PI City, State, Zip: Durham, NC 27708-0129
  - \* PI E-mail Address: awb@cs.duke.edu
  - \* PI URL Home Page: <http://www.cs.duke.edu/~awb>
  - \* Grant Title: Dialogue Theory for Virtual Environments
  - \* Grant/Contract Number: N00014-94-1-0938
  - \* Mipr Number:
  - \* R&T Number: 3331005vei01
  - \* Period of Performance: 1 July 1994 - 30 June 1996
  - \* Today's Date: 30 Nov. 96
- 

## **2. Productivity Measures.**

- \* Number of refereed papers submitted not yet published: 1
  - \* Number of refereed papers published: 10
  - \* Number of unrefereed reports and articles: 1
  - \* Number of books or parts thereof submitted but not published: 0
  - \* Number of books or parts thereof published: 0
  - \* Number of project presentations: 2
  - \* Number of patents filed but not yet granted: 0
  - \* Number of patents granted and software copyrights: 0
  - \* Number of graduate students supported  $\geq 25\%$  of full time: 0
  - \* Number of post-docs supported  $\geq 25\%$  of full time: 2
  - \* Number of minorities supported: 0
- 

## **3. Summary of Objectives and Approach.**

1. Our project has developed a theory of dialogue that enables a machine to cooperate with a human in the solution of a problem. Specifically, the machine proceeds to prove the top level goal that represents the solution to the problem. If it finds subgoals in its proof that it cannot solve, it resorts to an interaction with the user to attempt to obtain the needed information to finish the proof. Carried to its natural conclusion, this strategy becomes a theory of dialogue; all interactions are initiated to fill in gaps in uncompleted proofs.

This theory was implemented in a circuit repair system that could help a user diagnose and repair a failure in an electric circuit. The system could interact with the user with voice and was perfected and tested at length.

The current project aims to extend the mechanism to handle multimedia interactions with the user. Specifically we have been implementing a multimedia grammatical system that can handle a variety of modes such as voice, graphics

entities, displayed text, artificial sounds, and haptic devices. With this system, our dialogue mechanisms can proceed as before but the interactions can utilize all of these communicative modes.

The approaches to the research involve both developing a theoretical model and studying its properties and implementing the ideas in a voice-graphics interactive dialogue machine. The particular system currently being prototyped is a tutor for teaching computer programming.

---

#### **4. Detailed Summary of Technical Progress.**

1. The system has been coded and built into a programming tutorial system so that it can be examined in action and tested on real users.
2. The first such test occurred in September 1995 when a group of Duke programming students received one lesson in their regular programming course using the multimedia interactive system. Our project collected a complete trace of their behaviors in using the system and was able to find a variety of improvements needed for it to be efficient for such users. This led to a redesign of the main architecture and substantial new features.
3. One of the issues that arose is which media to use and what form should the message take when delivering a communication to the user. One could use speech with a blinking arrow aimed at the target object, one could use text highlighting and a typed example on the screen for the user to see, or any of many other formats. We decided to code an adaptive algorithm that randomly selects an output form, collects data on the quality of the user's response, and then increases its use of the forms that seem to work best. This learning algorithm was developed and has been installed in our system.
4. A second system test was conducted in September 1996. In this set of tests we hoped to evaluate a new temperature based dialogue algorithm, as well as the effectiveness of an animated face that we hoped would help to engage the user. Again, we asked students from the introductory Computer Science class to use our program in place of one of their regular weekly lab sessions. 13 students used the system, some with facial animations controlled by a researcher, the rest without faces, all using the temperature based dialogue model.

---

#### **5. Transitions and DOD Interactions.**

1. Our group has had negotiations with the Army Research Office regarding these technologies and the result of these has been the initiation of an industrial project to create voice and graphics applications for the U. S. Army. This project has been undertaken by the Research Triangle Institute in Research Triangle Park, North Carolina with the first task being the creation of a voice-virtual environment tutor and trainer for tank repair and maintenance. The work began in February of 1995 funded at the level of approximately \$600,000 for the first year. It resulted in a prototype system being created which has been demonstrated to the sponsor.

## 6. Software and Hardware Prototypes.

1. Prototype Name: The Duke Programming Tutor
  - Type: Software tutoring system
  - URL: <http://www.cs.duke.edu/~msf/voicelab/tutor/>
  - Availability: Can be made available if required.
  - Description: This system enables a programming student to type in a test program and then carry out a voice-graphics dialogue to discover the errors in the program and to fix them. The system currently assumes the program is written in the Pascal language and it needs a prototype correct program to compare with the student's code.
  - Demonstration Examples: A typical session begins with the student typing in a PASCAL homework program. The student then interacts with the computer and fixes the bugs. Shown below is a transcript of an actual dialogue.

User: Is anything wrong with my program?

Computer: There is something wrong with the write line statement.

User: I need more help.

Computer: There is something wrong with this string.

User: I do not understand.

Computer: Please delete the character space before the left parenthesis in the 2nd indented line.

User: Okay.

Computer: There is something wrong with the read line statement.

User: Okay.

Computer: There is something wrong with the if statement.

User: Where is the problem?

Computer: There is something wrong with the then part of the if statement.

User: I need more help.

Computer: The write line statement is still not correct.

User: Okay.

Computer: Continue working on your program.

User: Explain write line.

Computer: The write line statement is used when you want to output some value onto the screen during the execution of the program.

User: Okay.

Computer: To use the write line statement, you have to type in the following.

At this point, a window appears on the screen giving the syntax of the write line statement and explaining it's arguments.

User (to experimenter) : Is it okay now?

Experimenter : Continue asking it.

User: Is anything wrong with my program?

Computer: The write line statement is still not correct.

User: I need more help.

Computer: There is something wrong with the string of the write line statement.

User: I do not understand.

Computer: There is something wrong with this string.

User: I need more help.  
 Computer: You need to insert a character 'r' here.  
 User: Okay.  
 Computer: Congratulations, your program is correct now.

End of Session.

---

## 7. List of Publications.

1. Greg Keim, Michael Fulkerson and Alan Biermann. Initiative in Tutorial Dialogue Systems. 1997 AAAI Spring Symposium on Computational Models for Mixed Initiative Interactions.
2. P. Auer, P.M. Long and A. Srinivasan. Approximating Hyper-rectangles: Learning and Pseudorandom Sets. Proceedings of the 29th Annual ACM Symposium on the Theory of Computation, 1997.
3. R.D. Barve and P.M. Long. On the Complexity of Learning from Drifting Distributions. Proceedings of the 1996 Conference on Computational Learning Theory.
4. A.W. Biermann, C.I. Guinn, M. Fulkerson, G. Keim, Z. Liang, D. Melamed, and K. Rajagopalan. Goal-Oriented Multimedia Dialogue with Variable Initiative. Submitted for journal publication, 1996.
5. A.W. Biermann and P.M. Long. The Composition of Messages in Speech-Graphics Interactive Systems. Proceedings of the 1996 International Symposium on Spoken Dialogue.
6. N. Cesa-Bianchi, P.M. Long and M.K. Warmuth. Worst-case Quadratic Loss Bounds for Prediction using Linear Functions and Gradient Descent. IEEE Transactions on Neural Networks, 7(3):604-619, 1996.
7. C.I. Guinn. The Role of Computer-Computer Dialogues in Human-Computer Dialogue System Development. Empirical Methods in Discourse Interpretation and Generation, Proceedings of the AAAI 1995 Spring Symposium.
8. C.I. Guinn. Mechanisms for Mixed-Initiative Human-Computer Collaborative Discourse. Proceedings of the 34th Annual Meeting of the Association for Computational Linguistics, 1996.
9. C.I. Guinn. Mechanisms for Dynamically Changing Initiative in Human-Computer Collaborative Discourse. Submitted for publication, 1996.
10. P.M. Long. Improved Bounds about On-line Learning of Smooth Functions of a Single Variable. Proceedings of the 1996 Workshop on Algorithmic Learning Theory. Invited to the special issue of Theoretical Computer Science for ALT'96.
11. P.M. Long, A.I. Natsev and J.S. Vitter. Text Compression via Alphabet Representation. Proceedings of the 1997 Data Compression Conference.
12. P.M. Long and L. Tan. PAC Learning Axis-Aligned Rectangles with Respect to Product Distributions from Multiple-instance Examples. Proceedings of the 1996 Conference on Computational Learning Theory. Invited to the special issue of Machine Learning for COLT'96.

---

## 8. Invited and Contributed Presentations.

1. Alan W. Biermann, "Dialogue Theory for Virtual Environments", Office of Naval Research Virtual Environment Workshop, Arlington Virginia, 21-24 March, 1995.

2. Alan W. Biermann, "Spoken Language Dialogue", Invited lecture, Spoken Human-Machine Dialogue Workshop, U. S. Army Training and Doctrine Command, Research Triangle Park, North Carolina, 30 May- 1 June 1995.
- 

#### **9. Honors, Prizes or Awards Received.**

1. Alan W. Biermann was made a Fellow of the AAAI in August, 1994.
  2. Alan W. Biermann is currently Chairman, National Research Council Study: Towards an Every-Citizen Interface to the National Information Infrastructure.
- 

#### **10. Project Personnel Promotions.**

---

#### **11. Project Staff.**

1. Name: Dr. Curry I. Guinn
    - Homepage
    - Position: Research Associate
    - Task: Theoretical work and programming
  2. Name: Dr. Phil Long
    - Homepage
    - Position: Research Associate
    - Task: Theoretical work and programming
- 

#### **12. Multimedia URL.**

1. EOYL FY95
  2. QUAD FY95
  3. EOYL FY94
  4. Dialogue Theory for Virtual Environments
  5. Quicktime Movie of a Real Debugging Session
- 

#### **13. Keywords.**

1. Voice Interactive Systems
  2. Multimedia Systems
  3. Human-Machine Interface
  4. Virtual Environments
-

#### **14. Business Office**

- \* Business Office Phone Number: 919-684-5442
  - \* Business Office Fax Number: 919-684-8377
  - \* Business Office Email: johnsonme@mail01.adm.duke.edu
- 

#### **15. Expenditures**

1. FY95: 100%
  2. FY94: 100%
- 

#### **16. Current and Former Students**

1. Name: Mr. Michael S. Fulkerson
    - Homepage
    - Position: Second Year Ph.D. Student
    - Nationality: United States
    - Task: Programmer and Researcher
    - Thesis:
    - Graduated:
    - Job:
  2. Name: Mr. Greg A. Keim
    - Homepage
    - Position: Third Year Ph.D. Student
    - Nationality: United States
    - Task: Programmer and Researcher
    - Thesis:
    - Graduated:
    - Job:
  3. Name: Mr. Krishnan Rajagopalan
    - Homepage
    - Position: Graduated M.S. Student
    - Nationality: India
    - Task: Programmer and Researcher
    - Thesis: Development of a Theory for Control in Human Computer Interactions
    - Graduated: May 1996
    - Job: Cambridge Technology Partners
-

## 17. Book Plans

---

## 18. Sabbatical Plans

---

## 19. Related Research

1. James F Allen, University of Rochester
2. BBN, The HARK recognizer
3. Barbara Grosz, Harvard
4. Kathleen McKeown, Columbia University Natural Language Group
5. Steven K. Feiner, Computer Graphics and User Interfaces Lab, Columbia University
6. Oregon Graduate Institute's Center for Spoken Language Understanding

### NewsGroups

1. comp.ai.nat-lang
  2. comp.human-factors
  3. comp.ai.nlang-know-rep
  4. comp.multimedia
  5. comp.speech
  6. comp.ai.edu
  7. comp.ai.fuzzy
- 

## 20. History

1. We listed above that our project has passed its results (including some actual code) to the Research Triangle Institute, Research Triangle Park, North Carolina. This has resulted in the ARO sponsored project described there.

As a second example, our project created a voice interactive word processing system in the mid 1980s called VIPX. This system has been undertaken as the prototype for a Kurzweil AI, Inc. product development project in Waltham, Massachusetts. It is funded by NIST. The development is going forward at this time and has already been demonstrated in a prototype form.

2. In 1979, our project ran an experiment to test a typed natural language programming system. Students were asked to create programs using a special subset of English and using a traditional programming language and their performances using the two approaches were compared. One of the most noticeable aspects of student behaviors in the study was that the mere entering of the program was a major impediment to using the natural language system. This led to the addition of a voice interface to that system and, in later years, to a series of voice interactive systems.

In the mid 1980s, dialogue theory, as described by James Allen, Barbara Grosz, Candy Sidner, our project, and several others, came into existence. This theory emphasized the idea of subdialogues as a major construct of dialogues and



proposed ways to decompose interactions into such subunits. Our Circuit Fixit Shop Project between 1988 and 1991 was an implementation that tested many of these ideas. We chose PROLOG for our representation of knowledge and created our missing axiom theory for driving the interaction. In later work, we have been trying to formalize this theory to the point that its properties can be investigated in more systematic ways.

Our next step was to turn to another idea that had been set aside for a long time. We had, for most part, ignored other interaction modalities besides typed and spoken natural language. We decided to create a multimedia grammar to do the translation between the internal logical language and the various external modes available to communication. This model includes a complexity feature that measures the desirability of the external form when it is generated and helps the system prefer attractive and efficient forms of expression.

---